CLAIMS

What is claimed is:

1. A method comprising:

extracting URLs from electronic communication; and

analyzing the URLs extracted to determine whether the electronic communication is of a first predetermined category.

2. The method of claim 1, wherein extracting the URLs comprises extracting at least one of a hostname, a domain name, a subsection of a domain relative link, and an Internet Protocol (IP) address from the electronic communication.

3. The method of claim 1, further comprising performing a predetermined operation on the electronic communication if the electronic communication is determined to be of the first predetermined category.

4. The method of claim 1, wherein analyzing the URLs comprises:

generating one or more signatures based on the URLs extracted;

selecting one or more of the one or more signatures generated; and

comparing the selected signatures against a plurality of predetermined signatures generated from a plurality of known electronic communications of the first predetermined category.

5. The method of claim 4, wherein generating the one or more signatures further comprises using a length of the electronic communication to generate the one or more signatures.

6. The method of claim 4, wherein generating the one or more signatures further comprises using the extracted URLs as the one or more signatures.

7. The method of claim 4, wherein generating the one or more signatures further comprises generating the one or more signatures based on at least one of a protocol, a hostname, a domain name, a subsection of a domain relative link, and an Internet Protocol (IP) address from the electronic communication.

8. The method of claim 4, further comprising classifying the electronic communication to be of the first predetermined category if one of the selected signatures matches one of the plurality of predetermined signatures.

9. The method of claim 4, wherein the plurality of predetermined signatures is derived from a plurality of electronic documents reported via a collaborative submission mechanism.

10. A machine-accessible medium that provides instructions that, if executed by a processor, will cause the processor to perform operations comprising:

generating one or more signatures of electronic communication based on URLs in the electronic communication; and

determining whether the electronic communication is of a first predetermined category using the one or more signatures generated.

11. The machine-accessible medium of claim 10, wherein determining whether the electronic communication is of the first predetermined category comprises:

selecting one or more of the one or more signatures generated based on a plurality of predetermined criteria;

comparing the selected signatures against a plurality of predetermined signatures; and

classifying the electronic communication to be of the first predetermined category if one of the selected signatures matches one of the plurality of predetermined signatures.

12. The machine-accessible medium of claim 11, wherein selecting one or more of the one or more signatures generated comprises selecting a signature if the signature represents a domain that was registered within a predetermined period of time.

13. The machine-accessible medium of claim 11, wherein selecting one or more of the one or more signatures generated comprises selecting signatures representing one or more of a protocol, a hostname, a domain name, and a subsection of a domain relative link having a predetermined string of letters.

14. The machine-accessible medium of claim 10, wherein the operations further comprise extracting the URLs from the electronic communication.

15. A system comprising:

a plurality of databases to store a plurality of predetermined signatures of a plurality of known electronic communications of a first predetermined category; and

a server, coupled to the plurality of databases, including:

a memory device to store a plurality of instructions; and

a processor, coupled to the memory device, to retrieve the plurality of instructions from the memory device and to perform operations in response to the plurality of instructions, the operations comprising:

extracting URLs from electronic communication to generate one or more signatures; and

comparing one or more of the one or more signatures generated against the plurality of predetermined signatures stored in the plurality of databases to determine whether the electronic communication is of the first predetermined category.

16. The system of claim 15, wherein the URLs comprises at least one of a hostname, a domain name, a subsection of a domain relative link, and an Internet Protocol (IP) address.

17. The system of claim 15, wherein the operations further comprise selecting the one or more of the plurality of signatures based on a plurality of predetermined criteria.

18. The system of claim 15, wherein the operations further comprise performing a predetermined operation on the electronic communication if the electronic communication is determined to be of the first predetermined category.

19. The system of claim 15, further comprising a database, coupled to the server, to store a plurality of reports from which the plurality of predetermined signatures are generated.

20. The system of claim 15, wherein the plurality of databases are in a remote location from the server.